AMENDMENTS TO THE SPECIFICATION

On page 1, please replace the paragraph immediately below the title, with the following rewritten paragraph:

-- This application is a continuation of ~~copending~~ U.S. Application No. 10/074,279, filed February 12, 2002, now U.S. Patent No. 6,646,184, which ~~claims benefit~~ is a continuation of U.S. Application No. 09/538,414, filed March 29, 2000, now U.S. Patent No. 6,346,655, which claims the benefit of U.S. Provisional Application No. 60/304,177, filed March 31, 1999, and U.S. Provisional Application No. 60/287,549, filed February 11, 2000, all of which are incorporated herein by reference in their entirety.--

Please replace the first full paragraph on page 5, lines 3-25, with the following rewritten paragraph.

--In its broadest sense, the term "substantially similar", when used herein with respect to a nucleic acid molecule, means a nucleic acid molecule corresponding to a reference nucleotide sequence, wherein the corresponding nucleic acid molecule encodes a polypeptide having substantially the same structure and function as the polypeptide encoded by the reference nucleotide sequence, e.g. where only changes in amino acids not affecting the polypeptide function occur. Desirably the substantially similar nucleic acid molecule encodes the polypeptide encoded by the reference nucleotide sequence. The term "substantially similar" is specifically intended to include nucleic acid molecules wherein the sequence has been modified to optimize expression in particular cells, e.g. in plant cells. The percentage of identity between the substantially similar nucleic acid molecule and the reference nucleotide sequence desirably is at least 45%, more desirably at least 65%, more desirably at least 75%, preferably at least 85%, more preferably at least 90%, still more preferably at least 95%, yet still more preferably at least 99%. Preferably, the percentage of identity exists over a region of the sequences that is at least about 50 residues in length, more

preferably over a region of at least about 100 residues, and most preferably the sequences are substantially similar over at least about 150 residues. In a most preferred embodiment, the sequences are substantially similar over the entire length of the coding regions. Sequence comparisons may be carried out using a Smith-Waterman sequence alignment algorithm and as described in more detail below (see e.g. Waterman, M.S. Introduction to Computational Biology: Maps, sequences and genomes. Chapman & Hall. London: 1995. ISBN 0-412-99391-0~~, or at http://www-hto.usc.edu/software/seqaln/index.html~~). The local S program, version 1.16, is used with following parameters: match: 1, mismatch penalty: 0.33, open-gap penalty: 2, extended-gap penalty: 2.--

Please replace the paragraph on page 8, beginning at line 11, and continuing to page 9, through line 2, with the following rewritten paragraph.

--One example of an algorithm that is suitable for determining percent sequence identity and sequence similarity is the BLAST algorithm, which is described in Altschul *et al., J. Mol. Biol.* 215: 403-410 (1990). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. ~~(http://www.ncbi.nlm.nih.gov/).~~ This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, 1990). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when the cumulative alignment score falls off by the quantity X from its maximum achieved value, the cumulative score goes to zero or below due to the accumulation of one or more negative-scoring residue alignments, or the end of either sequence is reached.

The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, a cutoff of 100, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix *(see* Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89: 10915 (1989)).--